BOEING SCIENTIFIC RESEARCH LABORATORIES DOCUMENT D1-82-0762

BOEING

SCIENTIFIC RESEARCH LABORATORIES

A New Family of Life Distributions

Z. W. Birnbaum

Sam C. Saunders

MATHEMATICS RESEARCH OCTOBER 1968

19

D1-82-0762

A NEW FAMILY OF LIFE DISTRIBUTIONS

by

Z. W. Birnbaum
University of Washington

and

Sam C. Saunders

Mathematical Note No. 576

Mathematics Research Laboratory

BOEING SCIENTIFIC RESEARCH LABORATORIES

October 1968

SUMMARY

A new two parameter family of life length distributions is presented which is derived from a model for fatigue. This derivation follows from considerations of renewal theory for the number of cycles needed to force a fatigue crack extension to exceed a critical value. Some closure properties of this family are given and some comparisons made with other families such as the lognormal which have been previously used in fatigue studies.

2. A MODEL FOR THE DISTRIBUTION OF LIFE

We propose a basic framework and notation which is similar to that used previously in [3]. We consider only standardized material specimens which are subjected to fluctuating stresses by a periodic loading. By a load (or load oscillation) we mean a continuous unimodal function on the unit interval, the value of which at any time gives the stress imposed by the deflection of the specimen. Let $\ell_1, \ell_2, \ldots$ be the sequence of loads which are to be applied at each oscillation so that at the i^{th} oscillation load ℓ_i is imposed. We suppose that the loading is *cyclic* in the sense that for some $m > 1$ and all $i=1,\ldots,m$

$$\ell_{jm+i} = \ell_{km+i} \quad \text{for all} \quad j \neq k \tag{2.1.1}$$

and the loading is continuous so that for all $i=1,2,\ldots$

$$\ell_{i+1}(0) = \ell_i(1). \tag{2.1.2}$$

Hence the $(j+1)^{st}$ *cycle* is the loading $(\ell_{jm+1},\ldots,\ell_{jm+m})$.

We assume that fatigue failure is due to the initiation, growth and ultimate extension of a dominant crack. At each oscillation this crack is extended by some amount which is a random function due to the variation in the material, the magnitude of the imposed stress and a certain number of the prior loads and perhaps the actual crack extensions caused by the prior loads in that cycle.

Thus we now make our first assumption

1° The incremental crack extension X_i following the application of the i^{th} oscillation is a random variable with a distribution which depends upon all and only the loads and actual crack extensions which have preceded it in that cycle.

1. INTRODUCTION

It is well known that for the amount of fatigue data which can usually be obtained almost any two dimensional parametric family of distributions can be made to fit reasonably well. In fact, in the region of central tendency the lognormal, the Weibull, the Gamma, etc., can all be fitted by parametric estimation and because of the relatively small sample sizes hardly any can be rejected by, say a Chi-square Goodness of Fit test. However, when it becomes a question of predicting the "safe life" say the one thousandth percentile, there is a wide discrepancy between these models.

For this reason a family of distributions which is obtained from considerations of the basic characteristics of the fatigue process should be more persuasive in its implications than any ad hoc family chosen for extraneous reasons. In this paper we derive, using some elementary renewal theory, a two parameter family of nonnegative random variables as an idealization of the number of cycles necessary to force a fatigue crack to grow to a critical value. We then examine some of its relevant properties.

This assumption can be plausibly held for ground-air-ground cycles in aeronautical fatigue studies and other such applications.

The crack extension during the $(j+1)^{st}$ cycle is

$$Y_{j+1} = X_{jm+1} + \cdots + X_{jm+m} \quad \text{for} \quad j=0,1,\ldots,$$

where X_{jm+i} is the (possibly microscopic) crack extension following the load ℓ_i applied in the i^{th} oscillation of the $(j+1)^{st}$ cycle.

It follows from Assumption 1°, regardless of how much dependence exists between the successive random extensions per oscillation in each cycle, that the random total crack extensions per cycle are independent.

Thus we could formally make a second assumption

2° The total crack extension Y_j due to the j^{th} cycle is a random variable with mean μ and variance σ^2 for all $j=1,2,\ldots$.

Our notation will be

$$W_n = \sum_{j=1}^{n} Y_j,$$

which has distribution function

$$H_n(w) = P[W_n \le w], \quad \text{for} \quad n=1,2,\ldots .$$

It follows from elementary probability considerations, see p. 189, [6], that the distribution of C, the number of such cycles until failure, in the case failure is defined as the crack length exceeding some fixed critical length ω for the first time, is

$$P[C \leq n] = 1 - H_n(\omega). \tag{2.2}$$

Now

$$P[C \leq n] = P\left[\sum_{i=1}^{n} \frac{Y_i - \mu}{\sigma\sqrt{n}} > \frac{\omega - n\mu}{\sigma\sqrt{n}}\right] \tag{2.3}$$

and since the Y_j are independent and identically distributed random variables, using the central limit theorem for n large, Equation (2.3) can be approximated by the normal distribution. Let $\mathfrak{N}$ be the distribution function of the standard normal variate with zero mean and unit variance defined for $-\infty < y < \infty$ by

$$\mathfrak{N}(y) = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt. \tag{2.4}$$

We now make explicit this assumption of equality

3° the distribution of C from (2.2) is

$$P[C \leq n] = 1 - \mathfrak{N}\left(\frac{\omega - n\mu}{\sqrt{n}\,\sigma}\right) = \mathfrak{N}\left(\frac{\sqrt{n}\,\mu}{\sigma} - \frac{\omega}{\sqrt{n}\,\sigma}\right).$$

We are cognizant of the fact that Assumption 3° takes as exact, for the applications intended, an approximate equality. Such an assumption can only be justified on physical grounds, not mathematical.

Firstly, if we suppose that we are dealing with a long cycle of oscillations which is as complex as the ground-air-ground cycle in aeronautics it might be reasonable to assume that each cycle itself consists of a large number of distinct phases of loading. Even though the total crack extension per cycle is the sum of random variables which are not necessarily either independent or identically distributed,

the number of summands might be sufficiently large to make it reasonable to assume that the total crack extension per cycle, namely Y_j, is itself approximately normal.

Secondly, because in the case mentioned above, fatigue is measured in the thousands of cycles and in many other applications in the millions of cycles, it is felt that such large numbers of random variables make 3° an eminently practical assumption especially in view of its simple analytic form.

This distributional form, parameterized differently, has been previously obtained by Freudenthal and Shinozuka in [7]. In that report, which was unpublished, only a heuristic derivation based on engineering considerations was presented and then ad hoc fitting procedures were used to substantiate the validity of that model with several sets of fatigue data.

As a possible alternative one could postulate that the distribution of the crack extension might be different for the earlier cycles than it would be for the later ones. One of the first suppositions might be to make the distribution of the extension per cycle depend upon the size of the crack at the start of the cycle. One such assumption and the resulting distribution of the total crack length at the end of n cycles is presented in Section 4. We do not make that assumption here but instead, for reasons of simplicity, proceed with the study of the implications of the one above.

We write

$$\alpha = \frac{\sigma}{\sqrt{\mu\omega}}, \qquad \beta = \frac{\omega}{\mu} \tag{2.5}$$

and replace n by the nonnegative real variable $t > 0$. If we now denote the continuous extension of the discrete random variable C by T, a continuous nonnegative random variable, then it follows by 3° and (2.5) that T has the life distribution

$$F(t:\alpha,\beta) = \mathfrak{N}[\frac{1}{\alpha}\,\xi(t/\beta)] \quad \text{for} \quad t > 0 \tag{2.6}$$

where

$$\alpha > 0, \qquad \beta > 0$$

and

$$\xi(t) = t^{\frac{1}{2}} - t^{-\frac{1}{2}}. \tag{2.7}$$

This two-parameter family of distributions is a plausible model for the distribution of fatigue life. The set of all life distributions of the form (2.6) for $\alpha,\beta > 0$ will be denoted by $\mathcal{F}$ and most of this paper will be devoted to the study of its properties. We shall also refer to the law which has the distribution (2.6) with the somewhat shorter notation $F(\alpha,\beta)$.

Of course, there are motivational derivations for other distributions as well. The Weibull distribution, which is well known for its applications as a life length distribution and for fatigue life in particular, is obtained as a special case of the extreme value distributions, see p. 302, [8]. The Gamma family has also been obtained as a distribution of life by utilizing a model of a bundle of strands which are supporting a tensile load, see [4].

It is instructive to make a comparison between the derivation of the family of distributions $\mathcal{F}$ and an appropriate adaptation of

the classical heuristic argument, found for example p. 219, Cramér [5], as it might be used to obtain the lognormal distribution of the time until failure in fatigue. This argument is presented, along with derivations of other distributions, by Parzen in [9]. We should also mention that damage accumulation due to fatigue was first treated as a renewal problem in that paper.

3. SOME PROPERTIES OF $\mathcal{G}$

Let T have the distribution defined in (2.6). Note that T is a two-parameter random variable with β as a location parameter since it is the median of the distribution. (We show later that β is neither the mode nor the mean.) Notice also that α is a shape parameter and β a scale parameter. As we have seen $\frac{1}{\alpha}\xi(T/\beta)$ is a standard normal random variable with mean zero and unit variance. If we let X be $\mathfrak{N}(0, \frac{\alpha^2}{4})$ we see that, in distribution,

$$2X = \xi(T/\beta). \tag{3.1}$$

If we define the function ψ by

$$\psi(x) = \xi^{-1}(2x) \quad \text{for all real } x, \tag{3.1.5}$$

then

$$T = \beta\psi(X). \tag{3.2}$$

From elementary algebra we find that

$$\psi(x) = [\rho(x)]^2 \tag{3.3}$$

where

$$\rho(x) = x + \sqrt{x^2+1}\ . \tag{3.4}$$

So by (3.2)

$$T = \beta[1 + 2X^2 + 2X\sqrt{1+X^2}\ ] \tag{3.5}$$

where X is $\mathfrak{N}(0, \frac{\alpha^2}{4})$. Hence we have immediately

$$E(T) = \beta(1 + \frac{\alpha^2}{2}) \tag{3.5.1}$$

$$E(T^2) = \beta^2(1 + 2\alpha^2 + \frac{3\alpha^4}{2})$$

$$\mathrm{var}(T) = (\alpha\beta)^2(1 + \frac{5\alpha^2}{4}) \qquad (3.5.2)$$

and we note that for fixed α the variance of T increases as the scale parameter (median) β increases. This is not true for the lognormal distribution but empirical evidence shows that it is true for the fatigue lives themselves.

By noting that

$$\frac{1}{\rho(x)} = -x + \sqrt{1+x^2} = \rho(-x) \qquad (3.5.3)$$

we see that, whenever $-X$ has the same distribution as X, we have by (3.3), in distribution,

$$\frac{1}{\rho(X)} = \rho(X) \quad \text{and} \quad \frac{1}{\psi(X)} = \psi(X). \qquad (3.6)$$

Thus there follows immediately from (3.2) the

Theorem 3.1. If T has a fatigue life distribution $F(\alpha,\beta)$, in $\mathcal{F}$ then $\frac{1}{T}$ has a distribution in $\mathcal{F}$ given by $F(\alpha, \frac{1}{\beta})$. Moreover, for any real $a > 0$, the random variable aT has a distribution in $\mathcal{F}$ given by $F(\alpha,a\beta)$.

It is known that every random variable with distribution defined by (2.2) for which the Y_j are nonnegative and have densities which are Pólya frequency functions of order 2, has an increasing failure rate, see [1]. While our random variable T does not have this property, its average failure rate is nearly nondecreasing. Specifically, it has the properties of our subsequent remarks (2.2) and (2.3). The class of failure rates which are nondecreasing on the average has been studied in [2] and was shown there to have some closure properties which are physically plausible.

<u>Remark 3.2</u>. T has an average failure rate which approaches a positive constant.

<u>Proof</u>. Without loss of generality, take $\alpha = \beta = 1$ and define

$$Q(t) = -\ell n\{1 - \mathfrak{N}[\xi(t)]\}, \qquad h(t) = \frac{Q(t)}{t} .$$

Writing Mill's ratio

$$M(\xi) = \xi[1 - \mathfrak{N}(\xi)]/\mathfrak{N}'(\xi)$$

then applying L'Hospital's rule to $h(t)$ for $t \to \infty$ and using (4.1.1) of the next section we obtain

$$h(t) \simeq \frac{\mathfrak{N}'[\xi(t)]}{1-\mathfrak{N}[\xi(t)]}\,\xi'(t) = \frac{(1-t^{-2})}{2M[\xi(t)]} .$$

We know that as $\xi \to \infty$ we have

$$M(\xi) \simeq 1 - \frac{1}{\xi^2} + \frac{3}{\xi^4} + O(\xi^{-5}).$$

Hence $\lim_{t\to\infty} h(t) = \frac{1}{2}$ but $Q(1) = \ell n\ 2 > \frac{1}{2}$. This proves our contention that h does not always increase but approaches a positive constant.||

<u>Remark 3.3</u>. Actual numerical calculation shows that the average failure rate h decreases slowly for $t > 1.64$, as illustrated in Figure 1.

Surprisingly enough we regard this particular feature of the distribution as a virtue since it is in agreement with the observed facts. Typical fatigue data, such as given in [7], show that we cannot assume that the failure rate is always increasing. Such data can be explained as being the mixture of distributions each with increasing failure rate such that the combined failure rate decreases for a time.

However, almost any model which is the mixture of distributions, even two Weibull distributions with the same shape parameter but different scale parameters, would necessarily be so cumbersome mathematically that the statistical problems which arise would be difficult analytically.

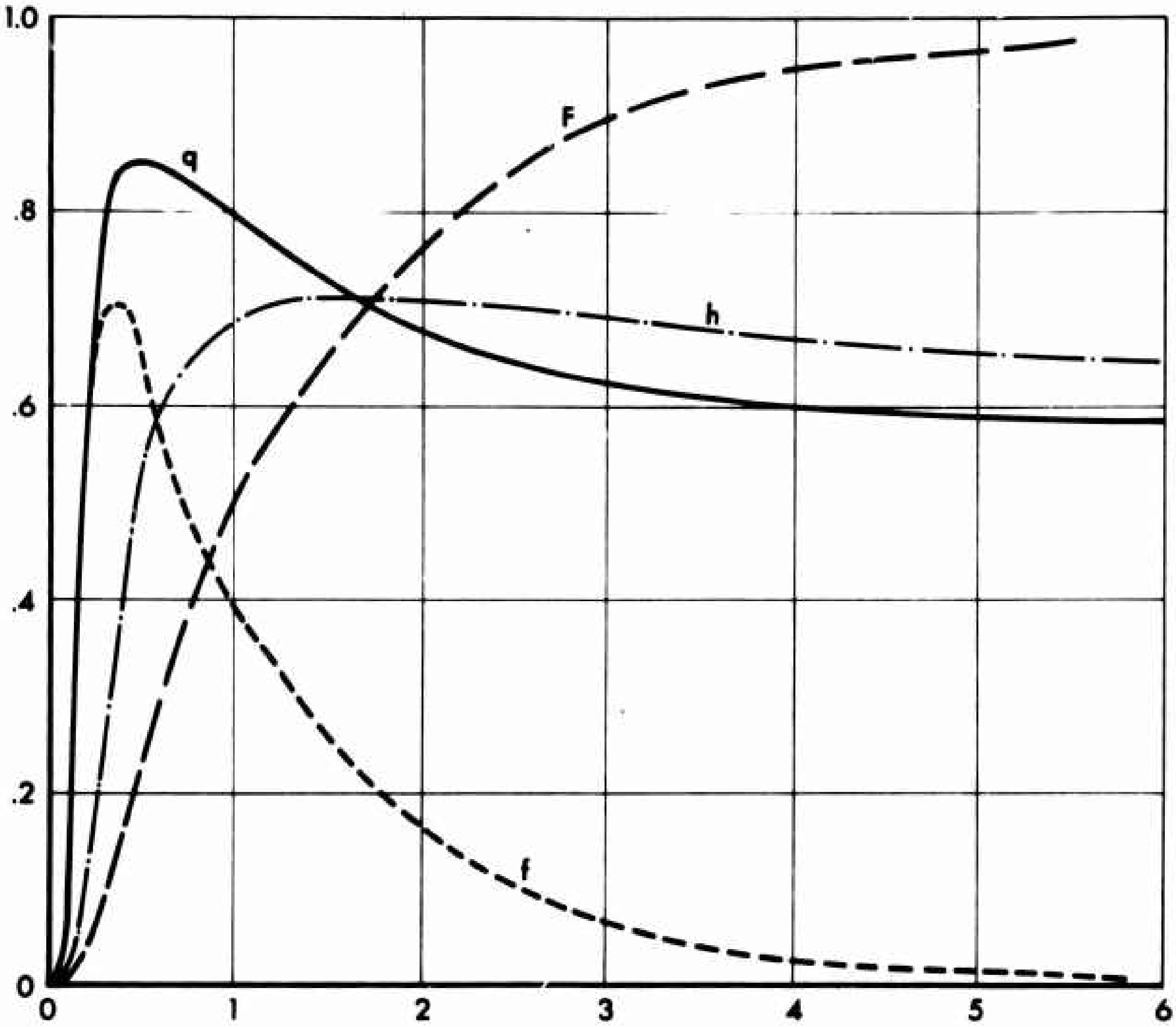

Figure 1

Graphs of the density f, the distribution F, the hazard rate q, and the average hazard rate h where for $t > 0$

$$f(t) = \frac{t^{\frac{1}{2}}+t^{-\frac{1}{2}}}{2t\sqrt{2\pi}} \exp\left\{-\frac{t}{2} - \frac{1}{2t} + 1\right\}$$

$$F(t) = \int_0^t f(x)dx, \quad q(t) = \frac{f(t)}{1-F(t)}, \quad h(t) = \frac{1}{t}\int_0^t q(x)dx.$$

4. A SECOND MODEL FOR THE DISTRIBUTION OF LIFE

In this section we replace the Assumption 2° by the alternative one

2' The crack extension Y_{n+1} during the $(n+1)^{st}$ cycle, given that the total crack length was s at the start of the cycle, is a normal random variable with mean $\mu + \delta s$, for some constant $\delta \geq 0$, and variance σ^2 for each $n=1,2,\ldots$.

By formula (2.2), to find the distribution of C it is necessary and sufficient that we find the distribution of W_n, called H_n.

By assumption

$$P[Y_{n+1} \leq y \mid W_n = s] = \mathfrak{N}\left(\frac{y-\mu-\delta s}{\sigma}\right)$$

and for $n=1$

$$H_1(y) = \mathfrak{N}\left(\frac{y-\mu}{\sigma}\right) \qquad -\infty < y < \infty. \tag{4.1}$$

Now by definition, setting $y = x - s$

$$H_{n+1}(x) = \int_{-\infty}^{\infty} \mathfrak{N}\left(\frac{y-\mu-\delta s}{\sigma}\right) dH_n(s). \tag{4.2}$$

We can now prove the

<u>Theorem 4.1</u>. The total crack length W_n, at the end of the n^{th} cycle, is normal with mean μ_n and variance σ_n^2 where

$$\mu_n = \frac{\mu}{\delta}\left[(1+\delta)^n - 1\right], \qquad \sigma_n^2 = \sigma^2 \frac{(1+\delta)^{2n}-1}{(1+\delta)^2-1}. \tag{4.3}$$

Proof by induction. The statement is true for $n=1$. Assume it true for n, then from (4.2)

$$H_{n+1}(x) = \int_{-\infty}^{\infty} \mathfrak{N}\left(\frac{u-s}{v}\right) d_s\, \mathfrak{N}\left(\frac{s-\mu_n}{\sigma_n}\right)$$

where

$$u = \frac{x-\mu}{1+\delta} \qquad v = \frac{\sigma}{1+\delta} .$$

But this we recognize as the usual convolution of two normal random variables and hence the result is

$$H_{n+1}(x) = \mathfrak{N}\left(\frac{u-\mu_n}{\sqrt{v^2+\sigma_n^2}}\right).$$

Hence by simplification we find that

$$\mu_{n+1} = \mu + (1+\delta)\mu_n, \qquad \sigma_{n+1}^2 = \sigma^2 + (1+\delta)^2\sigma_n^2$$

and one checks that formulas given do satisfy the recursion relations (4.3).||

Strictly speaking a proper distribution, analogous to that obtained in (2.6), which would be the continuous extension of the present case, cannot be generated since $P[C < \infty] < 1$.

For, one can see that

$$P[C \leq n] = \mathfrak{N}\left(\frac{\mu_n-\omega}{\sigma_n}\right)$$

and as $n \to \infty$ we have $\mu_n/\sigma_n \to \frac{\mu}{\delta\sigma} \sqrt{(1+\delta)^2-1} < \infty$.

Of course this would not be of practical significance, since we must restrict ourselves to situations where $\frac{\mu}{\sigma} > 3$, otherwise $2'$ would be physically unrealistic. Hence $\lim\limits_{n\to\infty} \frac{\mu_n}{\sigma_n} > 3\sqrt{2/\delta}$ for $0 < \delta < 1$, and $\lim \frac{\mu_n}{\sigma_n} > 3$ for any δ.

5. CONCLUSION

A derivation, based on plausible physical considerations, for a family of distributions is, by itself, not a conclusive argument that such a particular family should be used in life studies. No family, however reasonable its derivation, can be accepted for use in fatigue life studies until it is confronted with actual fatigue data obtained under various conditions and the distribution is shown to represent adequately the life lengths which are obtained.

In order to do this one must have the theory of estimation for this family completed. The derivation of parametric estimators and the ancillary computing formulas for this family will be presented in a latter study. Also further studies of the application of this distribution to the calculation of "safe life" will be made. Thus the confrontations of this family with actual data will be carried out, to provide the justification for this presentation.

BIBLIOGRAPHY

[1] Barlow, Richard E., and Proschan, Frank. Mathematical Theory of Reliability, John Wiley and Sons, New York, 1965.

[2] Birnbaum, Z. W., Esary, J. D., and Marshall, A. W. Stochastic Characterization of Wear-out for Components and Systems, *Ann. Math. Statist.*, 37, 816-825, 1966.

[3] Birnbaum, Z. W., and Saunders, S. C. A Probabilistic Interpretation of Miner's Rule, *SIAM J. Appl. Math.*, 16, 637-652, May 1968.

[4] Birnbaum, Z. W., and Saunders, S. C. A Statistical Model for Life-Length of Material, *J. Amer. Statist. Assoc.*, 53, 151-160, 1958.

[5] Cramér, Harold. Mathematical Methods of Statistics, Princeton University Press, 1951.

[6] Feller, William. An Introduction to Probability Theory and Its Applications, Volume II, John Wiley and Sons, New York, 1966.

[7] Freudenthal, A. M., and Shinozuka, M. Structural Safety under Conditions of Ultimate Load-failure and Fatigue, WADD Technical Report 61-77, October, 1961.

[8] Gumbel, E. J. Statistics of Extremes, Columbia University Press, New York, 1958.

[9] Parzen, Emanuel. On Models for the Probability of Fatigue Failure of a Structure, Report 245, North Atlantic Treaty Organization Advisory Group for Aeronautical Research and Development, 1959.